WHAT IS CLAIMED IS:

1. A device for automatically identifying the language of a digital text, comprising :

- means for prestoring first character strings that occur frequently anywhere respectively in words of a plurality of predetermined languages and characterize said predetermined languages,

- means for prestoring second character strings that are atypical anywhere respectively in words of said predetermined languages,

- means for analyzing words extracted from said digital text thereby constructing for each extracted word all character strings contained in said extracted word and having lengths lying between one character and the number of characters in said extracted word,

means for comparing character strings contained in extracted words to prestored character strings in order to determine scores associated with said predetermined languages,

- means for comparing each of all character strings contained in each said extracted word individually to said first and second prestored character strings of a determined language so that whenever a first character string is found in said extracted word a score associated with said determined language is increased by a first coefficient depending on the position of said first character string found in said extracted word and whenever a second character string is found in said extracted word said score is decreased by a respective

second coefficient that is associated with said found second character string and that increases as the probability of said found second character string in said determined language decreases, and

- means for comparing said scores for said text associated with said predetermined languages in order to determine the highest of said scores, which identifies the language of said text.

2. The device claimed in claim 1, wherein a first character string in an extracted word consists of one of the following character strings: a prefix, a pseudo-prefix, a suffix, a pseudo-suffix, an infix, a pseudo-infix.

3. The device claimed in claim 1, wherein said first coefficient of a first character string in said extracted word depends on the frequency of said character string in said determined language.

4. The device claimed in claim 1, wherein said first coefficient of a first character string in said extracted word depends on the length of said character string.

5. The device claimed in claim 1, wherein said first coefficient of a first character string in said extracted word is equal to :

PO (FR + LON),

where PO is a coefficient depending on the position of said first character string in said extracted word, FR is a coefficient depending on the frequency of said first

character string in a determined language, and LON is a coefficient depending on the length of said first character string.

6. The device claimed in claim 1, comprising comparator means for comparing each of said extracted words from said text with frequent words in said determined language and initially listed in storage means so that whenever a frequent word is found in said text said score for said determined language is increased only by a coefficient depending on the frequency of said extracted word in said determined language

7. The device claimed in claim 1, comprising comparator means for comparing each of said extracted words from said text with frequent words in said determined language and initially listed in storage means so that whenever a frequent word is found in said text said score for said determined language is increased only by a coefficient depending on the length of said frequent word.